

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日:

2004年6月24日(24.06.2004)

PCT

(10) 国际公布号:

WO 2004/052927 A1

(51) 国际分类⁷: C07K 14/47, C12N 15/12, A61K 38/17, A61P 17/14

(21) 国际申请号: PCT/CN2003/000956

(22) 国际申请日: 2003年11月13日(13.11.2003)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权: 02145253.9 2002年11月13日(13.11.2002) CN

(71) 申请人(对除美国以外的所有指定国): 中国科学院上海生命科学研究院(SHANGHAI INSTITUTES FOR BIOLOGICAL SCIENCES, CHINESE ACADEMY OF SCIENCES) [CN/CN]; 中国上海市岳阳路320号, Shanghai 200031 (CN)。

(72) 发明人: 及

(75) 发明人/申请人(仅对美国): 孔祥银(KONG, Xiangyin) [CN/CN]; 李善如(LI, Shanru) [CN/CN]; 胡兰毓(HU, Landian) [CN/CN]; 史耀舟(SHI, Yaozhou) [CN/CN]; 滕晓坤(TENG, Xiaokun) [CN/CN]; 中国上海市岳阳路320号, Shanghai 200031 (CN)。

(74) 代理人: 上海专利商标事务所(SHANGHAI PATENT & TRADEMARK LAW OFFICE); 中国上海市桂平路435号徐迅, Shanghai 200233 (CN)。

(81) 指定国(国家): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW

(84) 指定国(地区): ARIPO专利(BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), 欧亚专利(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), 欧洲专利(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI专利(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)

本国际公布:

— 包括国际检索报告。

所引用双字母代码和其它缩写符号, 请参考刊登在每期PCT公报期刊起始的“代码及缩写符号简要说明”。

(54) Title: A BALDNESS RELATED GENE AND THE POLYPEPTIDE ENCODED THEREBY, AND USES THEREOF

(54) 发明名称: 利用人和鼠Rhor基因及其编码产物诊断和治疗秃发的方法

(57) Abstract: The present invention disclosed a baldness-related gene and the polypeptide encoded thereby, and the recombinant method for producing the polypeptide. It also relates to the usage of said gene and polypeptide in the diagnosis and treatment of baldness.

(57) 摘要

本发明提供了一种秃发相关基因及其编码的多肽, 以及利用重组技术生产所述多肽的方法。本发明涉及了用所述基因和多肽在诊断和治疗秃发中的用途。

利用人和鼠 Rhor 基因及其编码产物诊断和治疗秃发的方法

技术领域

本发明属于生物技术和医学领域，具体地说，本发明涉及新的编码秃发相关蛋白-Rhor 的多核苷酸，以及此多核苷酸编码的多肽。本发明还涉及此多核苷酸和多肽的用途和制备。

背景技术

人类的先天性秃发、稀发常见于男性女性患者较为少见，虽然不影响人的生存质量，但是对于个人来讲由其是女性这种先天性的遗传疾病严重影响其生活质量，遗憾的是迄今为止没有可以从根本上治疗的方法。此外，搜集人类的先天性秃发家系存在较大的困难。

然而，目前还没有发现或分离出与秃发有关的基因。因此，本领域迫切需要开发新的新的与秃发有关的蛋白。

发明内容

本发明的目的是提供一种新的秃发相关蛋白-Rhor 蛋白以及其片段、类似物和衍生物。

本发明的另一目的是提供编码这些多肽的多核苷酸。

本发明的另一目的是提供生产这些多肽的方法以及该多肽和编码序列的用途。

在本发明的第一方面，提供新颖的分离出的 Rhor 多肽，它包含：具有 SEQ ID NO: 2 氨基酸序列的多肽、或其保守性变异多肽、或其活性片段、或其活性衍生物。较佳地，该多肽是具有 SEQ ID NO: 2 氨基酸序列的多肽。

在本发明的第二方面，提供编码分离的这些多肽的多核苷酸，该多核苷酸包含一核苷酸序列，该核苷酸序列与选自下组的一种核苷酸序列有至少 70%相同性：(a) 编码上述 Rhor 多肽的多核苷酸；和(b)与多核苷酸(a)互补的多核苷酸。较佳地，该多核苷酸编码具有 SEQ ID NO: 2 所示氨基酸序列的多肽。更佳地，该多核苷酸的序列具有 SEQ ID NO: 1 中 1-2484 位的序列。

在本发明的第三方面，提供了含有上述多核苷酸的载体，以及被该载体转化或转导的宿主细胞或者被上述多核苷酸直接转化或转导的宿主细胞。

在本发明的第四方面，提供了制备 Rhor 蛋白的方法，该方法包含：(a)在表达 Rhor 蛋白的条件下，培养上述被转化或转导的宿主细胞；(b)从培养物中分离出 Rhor 蛋白。

在本发明的第五方面，提供了与上述的 Rhor 多肽特异性结合的抗体。还提供

了可用于检测的核酸分子，它含有上述的多核苷酸中连续的 20-2484 个核苷酸。

在本发明的第六方面，提供了模拟、促进、拮抗 Rhor 多肽活性的化合物，以及抑制 Rhor 多肽的表达的化合物。还提供了筛选和/或制备这些化合物的方法。较佳地，该化合物是 Rhor 多肽的编码序列或其片段的反义序列。

在本发明的第七方面，提供了检测样品中是否存在 Rhor 蛋白的方法，它包括：将样品与 Rhor 蛋白的特异性抗体接触，观察是否形成抗体复合物，形成了抗体复合物就表示样品中存在 Rhor 蛋白。

在本发明的第八方面，提供了一种检测与 Rhor 多肽异常表达相关的疾病或疾病(如秃发)易感性的方法，该方法包括：检测编码所述多肽的核酸序列中是否存在突变。较佳地，所述的突变是缺失了 SEQ ID NO: 1 中 351—580 的核苷酸序列。此外，还提供了相应的试剂盒，它包括特异性扩增 Rhor 基因或转录本的引物。较佳地，所述试剂盒还含有与突变部位结合的探针。

在本发明的第九方面，提供了本发明多肽和编码序列的用途。例如本发明多肽可被用于筛选促进 Rhor 多肽活性的激动剂，或者筛选抑制 Rhor 多肽活性的拮抗剂、或者被用于肽指纹图谱鉴定。本发明的 Rhor 蛋白的编码序列或其片段，可被作为引物用于 PCR 扩增反应，或者作为探针用于杂交反应，或者用于制造基因芯片或微阵列。

在本发明的第十方面，提供了一种药物组合物，它含有安全有效量的本发明的 Rhor 多肽或其激动剂、拮抗剂以及药学上可接受的载体。这些药物组合物可治疗秃发等病症。

本发明的其它方面由于本文的技术的公开，对本领域的技术人员而言是显而易见的。

附图说明

图 1 是不同小鼠的 Rhor 基因片段的 PCR 扩增结果。其中泳道 1：无毛小鼠；2：稀毛小鼠；3 和 4：正常小鼠。

具体实施方式

本发明人经过深入而广泛的研究，首次分离和证明了一种新基因 Rhor 与秃发密切相关，该基因编码类似表皮生长因子受体的蛋白。Rhor 的突变或功能下降或缺失将直接导致秃发。在此基础上完成了本发明。

在本发明中，术语“Rhor 蛋白”、“Rhor 多肽”或“秃发相关蛋白 Rhor”可互换使用，都指具有秃发相关蛋白 Rhor 氨基酸序列(SEQ ID NO:2)的蛋白或多肽。它们包括含有或不含起始甲硫氨酸的 Rhor 蛋白。Rhor 也被称为“RHBDL5”。

如本文所用,“分离的”是指物质从其原始环境中分离出来(如果是天然的物质,原始环境即是天然环境)。如活体细胞内的天然状态下的多聚核苷酸和多肽是没有分离纯化的,但同样的多聚核苷酸或多肽如从天然状态中同存在的其他物质中分开,则为分离纯化的。

如本文所用,“分离的 Rhor 蛋白或多肽”是指 Rhor 多肽基本上不含天然与其相关的其它蛋白、脂类、糖类或其它物质。本领域的技术人员能用标准的蛋白质纯化技术纯化 Rhor 蛋白。基本上纯的多肽在非还原聚丙烯酰胺凝胶上能产生单一的主带。

本发明的多肽可以是重组多肽、天然多肽、合成多肽,优选重组多肽。本发明的多肽可以是天然纯化的产物,或是化学合成的产物,或使用重组技术从原核或真核宿主(例如,细菌、酵母、高等植物、昆虫和哺乳动物细胞)中产生。根据重组生产方案所用的宿主,本发明的多肽可以是糖基化的,或可以是非糖基化的。

本发明还包括 Rhor 蛋白的片段、衍生物和类似物。如本文所用,术语“片段”、“衍生物”和“类似物”是指基本上保持本发明的天然 Rhor 蛋白相同的生物学功能或活性的多肽。本发明的多肽片段、衍生物或类似物可以是(i)有一个或多个保守或非保守性氨基酸残基(优选保守性氨基酸残基)被取代的多肽,而这样的取代的氨基酸残基可以是也可以不是由遗传密码编码的,或(ii)在一个或多个氨基酸残基中具有取代基团的多肽,或(iii)成熟多肽与另一个化合物(比如延长多肽半衰期的化合物,例如聚乙二醇)融合所形成的多肽,或(iv)附加的氨基酸序列融合到此多肽序列而形成的多肽(如前导序列或分泌序列或用来纯化此多肽的序列或蛋白原序列,或与抗原 IgG 片段的形成的融合蛋白)。根据本文的教导,这些片段、衍生物和类似物属于本领域熟练技术人员公知的范围。

在本发明中,术语“Rhor 多肽”指具有 Rhor 蛋白活性的 SEQ ID NO: 2 序列的多肽。该术语还包括具有与 Rhor 蛋白相同功能的、SEQ ID NO: 2 序列的变异形式。这些变异形式包括(但不限于):若干个(通常为 1-50 个,较佳地 1-30 个,更佳地 1-20 个,最佳地 1-10 个)氨基酸的缺失、插入和/或取代,以及在 C 末端和/或 N 末端添加一个或数个(通常为 20 个以内,较佳地为 10 个以内,更佳地为 5 个以内)氨基酸。例如,在本领域中,用性能相近或相似的氨基酸进行取代时,通常不会改变蛋白质的功能。又比如,在 C 末端和/或 N 末端添加一个或数个氨基酸通常也不会改变蛋白质的功能。该术语还包括 Rhor 蛋白的活性片段和活性衍生物。

该多肽的变异形式包括:同源序列、保守性变异体、等位变异体、天然突变体、诱导突变体、在高或低的严紧度条件下能与 Rhor DNA 杂交的 DNA 所编码的蛋白、以及利用抗 Rhor 多肽的抗血清获得的多肽或蛋白。本发明还提供了其他多肽,如包含 Rhor 多肽或其片段的融合蛋白。除了几乎全长的多肽外,本发明还包括了 Rhor 多肽的可溶性片段。通常,该片段具有 Rhor 多肽序列的至少约 10 个连续氨基酸,

通常至少约 30 个连续氨基酸，较佳地至少约 50 个连续氨基酸，更佳地至少约 80 个连续氨基酸，最佳地至少约 100 个连续氨基酸。

发明还提供 Rhor 蛋白或多肽的类似物。这些类似物与天然 Rhor 多肽的差别可以是氨基酸序列上的差异，也可以是不影响序列的修饰形式上的差异，或者兼而有之。这些多肽包括天然或诱导的遗传变异体。诱导变异体可以通过各种技术得到，如通过辐射或暴露于诱变剂而产生随机诱变，还可通过定点诱变法或其他已知分子生物学技术。类似物还包括具有不同于天然 L-氨基酸的残基(如 D-氨基酸)的类似物，以及具有非天然存在的或合成的氨基酸(如 β 、 γ -氨基酸)的类似物。应理解，本发明的多肽并不限于上述列举的代表性的多肽。

修饰(通常不改变一级结构)形式包括：体内或体外的多肽的化学衍生形式如乙酰化或羧基化。修饰还包括糖基化，如那些在多肽的合成和加工中或进一步加工步骤中进行糖基化修饰而产生的多肽。这种修饰可以通过将多肽暴露于进行糖基化的酶(如哺乳动物的糖基化酶或去糖基化酶)而完成。修饰形式还包括具有磷酸化氨基酸残基(如磷酸酪氨酸，磷酸丝氨酸，磷酸苏氨酸)的序列。还包括被修饰从而提高了其抗蛋白水解性能或优化了溶解性能的多肽。

在本发明中，“Rhor 蛋白保守性变异多肽”指与 SEQ ID NO: 2 的氨基酸序列相比，有至多 10 个，较佳地至多 8 个，更佳地至多 5 个，最佳地至多 3 个氨基酸被性质相似或相近的氨基酸所替换而形成多肽。这些保守性变异多肽最好根据表 1 进行氨基酸替换而产生。

表 1

最初的残基	代表性的取代	优选的取代
Ala (A)	Val; Leu; Ile	Val
Arg (R)	Lys; Gln; Asn	Lys
Asn (N)	Gln; His; Lys; Arg	Gln
Asp (D)	Glu	Glu
Cys (C)	Ser	Ser
Gln (Q)	Asn	Asn
Glu (E)	Asp	Asp
Gly (G)	Pro; Ala	Ala
His (H)	Asn; Gln; Lys; Arg	Arg
Ile (I)	Leu; Val; Met; Ala; Phe	Leu
Leu (L)	Ile; Val; Met; Ala; Phe	Ile
Lys (K)	Arg; Gln; Asn	Arg
Met (M)	Leu; Phe; Ile	Leu
Phe (F)	Leu; Val; Ile; Ala; Tyr	Leu
Pro (P)	Ala	Ala

Ser (S)	Thr	Thr
Thr (T)	Ser	Ser
Trp (W)	Tyr; Phe	Tyr
Tyr (Y)	Trp; Phe; Thr; Ser	Phe
Val (V)	Ile; Leu; Met; Phe; Ala	Leu

本发明的多核苷酸可以是 DNA 形式或 RNA 形式。DNA 形式包括 cDNA、基因组 DNA 或人工合成的 DNA。DNA 可以是单链的或是双链的。DNA 可以是编码链或非编码链。编码成熟多肽的编码区序列可以与 SEQ ID NO:1 所示的编码区序列相同或者是简并的变异体。如本文所用,“简并的变异体”在本发明中是指编码具有 SEQ ID NO:2 所示氨基酸序列的蛋白质,但与 SEQ ID NO:1 所示的编码区序列有差别的核酸序列。

编码成熟多肽的多核苷酸包括:只编码成熟多肽的编码序列;成熟多肽的编码序列和各种附加编码序列;成熟多肽的编码序列(和任选的附加编码序列)以及非编码序列。

术语“编码多肽的多核苷酸”可以是包括编码此多肽的多核苷酸,也可以是还包括附加编码和/或非编码序列的多核苷酸。

本发明还涉及上述多核苷酸的变异体,其编码与本发明有相同的氨基酸序列的多肽或多肽的片段、类似物和衍生物。此多核苷酸的变异体可以是天然发生的等位变异体或非天然发生的变异体。这些核苷酸变异体包括取代变异体、缺失变异体和插入变异体。如本领域所知的,等位变异体是一个多核苷酸的替换形式,它可能是一个或多个核苷酸的取代、缺失或插入,但不会从实质上改变其编码的多肽的功能。

本发明还涉及与上述的序列杂交且两个序列之间具有至少 50%, 较佳地至少 70%, 更佳地至少 80% 相同性的多核苷酸。本发明特别涉及在严格条件下与本发明所述多核苷酸可杂交的多核苷酸。在本发明中,“严格条件”是指:(1)在较低离子强度和较高温度下的杂交和洗脱,如 $0.2 \times \text{SSC}$, 0.1% SDS, 60°C ; 或(2)杂交时加有变性剂,如 50%(v/v) 甲酰胺, 0.1% 小牛血清/0.1% Ficoll, 42°C 等; 或(3)仅在两条序列之间的相同性至少在 90% 以上,更好是 95% 以上时才发生杂交。并且,可杂交的多核苷酸编码的多肽与 SEQ ID NO:2 所示的成熟多肽有相同的生物学功能和活性。

本发明还涉及与上述的序列杂交的核酸片段。如本文所用,“核酸片段”的长度至少含 15 个核苷酸,较好是至少 30 个核苷酸,更好是至少 50 个核苷酸,最好是至少 100 个核苷酸以上。核酸片段可用于核酸的扩增技术(如 PCR)以确定和/或分离编码 Rhor 蛋白的多聚核苷酸。

本发明中的多肽和多核苷酸优选以分离的形式提供,更佳地被纯化至均质。

本发明的 Rhor 核苷酸全长序列或其片段通常可以用 PCR 扩增法、重组法或人工合成的方法获得。对于 PCR 扩增法,可根据本发明所公开的有关核苷酸序列,

尤其是开放阅读框序列来设计引物,并用市售的 cDNA 库或按本领域技术人员已知的常规方法所制备的 cDNA 库作为模板,扩增而得有关序列。当序列较长时,常常需要进行两次或多次 PCR 扩增,然后再将各次扩增出的片段按正确次序拼接在一起。

一旦获得了有关的序列,就可以用重组法来大批量地获得有关序列。这通常是将其克隆入载体,再转入细胞,然后通过常规方法从增殖后的宿主细胞中分离得到有关序列。

此外,还可用人工合成的方法来合成有关序列,尤其是片段长度较短时。通常,通过先合成多个小片段,然后再进行连接可获得序列很长的片段。

目前,已经可以完全通过化学合成来得到编码本发明蛋白(或其片段,或其衍生物)的 DNA 序列。然后可将该 DNA 序列引入本领域中已知的各种现有的 DNA 分子(或如载体)和细胞中。此外,还可通过化学合成将突变引入本发明蛋白序列中。

应用 PCR 技术扩增 DNA/RNA 的方法被优选用于获得本发明的基因。用于 PCR 的引物可根据本文所公开的本发明的序列信息适当地选择,并可用常规方法合成。可用常规方法如通过凝胶电泳分离和纯化扩增的 DNA/RNA 片段。

本发明也涉及包含本发明的多核苷酸的载体,以及用本发明的载体或 Rhor 蛋白编码序列经基因工程产生的宿主细胞,以及经重组技术产生本发明所述多肽的方法。

通过常规的重组 DNA 技术(Science, 1984; 224: 1431),可利用本发明的多聚核苷酸序列可用来表达或生产重组的 Rhor 多肽。一般来说有以下步骤:

(1).用本发明的编码 Rhor 多肽的多核苷酸(或变异体),或用含有该多核苷酸的重组表达载体转化或转导合适的宿主细胞;

(2).在合适的培养基中培养的宿主细胞;

(3).从培养基或细胞中分离、纯化蛋白质。

本发明中, Rhor 多核苷酸序列可插入到重组表达载体中。术语“重组表达载体”指本领域熟知的细菌质粒、噬菌体、酵母质粒、植物细胞病毒、哺乳动物细胞病毒如腺病毒、逆转录病毒或其他载体。在本发明中适用的载体包括但不限于:在细菌中表达的基于 T7 的表达载体(Rosenberg, et al. Gene, 1987, 56:125);在哺乳动物细胞中表达的 pMSXND 表达载体(Lee and Nathans, J Bio Chem. 263:3521,1988)和在昆虫细胞中表达的来源于杆状病毒的载体。总之,只要能在宿主体内复制和稳定,任何质粒和载体都可以用。表达载体的一个重要特征是通常含有复制起点、启动子、标记基因和翻译控制元件。

本领域的技术人员熟知的方法能用于构建含 Rhor 编码 DNA 序列和合适的转录/翻译控制信号的表达载体。这些方法包括体外重组 DNA 技术、DNA 合成技术、体内重组技术等(Sambrook, et al. Molecular Cloning, a Laboratory Manual, cold

Spring Harbor Laboratory, New York, 1989)。所述的 DNA 序列可有效连接到表达载体中的适当启动子上，以指导 mRNA 合成。这些启动子的代表性例子有：大肠杆菌的 lac 或 trp 启动子； λ 噬菌体 P_L 启动子；真核启动子包括 CMV 立即早期启动子、HSV 胸苷激酶启动子、早期和晚期 SV40 启动子、反转录病毒的 LTRs 和其他一些已知的可控制基因在原核或真核细胞或其病毒中表达的启动子。表达载体还包括翻译起始用的核糖体结合位点和转录终止子。

此外，表达载体优选地包含一个或多个选择性标记基因，以提供用于选择转化的宿主细胞的表型性状，如真核细胞培养用的二氢叶酸还原酶、新霉素抗性以及绿色荧光蛋白(GFP)，或用于大肠杆菌的四环素或氨苄青霉素抗性。

包含上述的适当 DNA 序列以及适当启动子或者控制序列的载体，可以用于转化适当的宿主细胞，以使其能够表达蛋白质。

宿主细胞可以是原核细胞，如细菌细胞；或是低等真核细胞，如酵母细胞；或是高等真核细胞，如哺乳动物细胞。代表性例子有：大肠杆菌，链霉菌属；鼠伤寒沙门氏菌的细菌细胞；真菌细胞如酵母；植物细胞；果蝇 S2 或 Sf9 的昆虫细胞；CHO、COS、293 细胞、或 Bowes 黑素瘤细胞的动物细胞等。

本发明的多核苷酸在高等真核细胞中表达时，如果在载体中插入增强子序列时将会使转录得到增强。增强子是 DNA 的顺式作用因子，通常大约有 10 到 300 个碱基对，作用于启动子以增强基因的转录。可举的例子包括在复制起始点晚期一侧的 100 到 270 个碱基对的 SV40 增强子、在复制起始点晚期一侧的多瘤增强子以及腺病毒增强子等。

本领域一般技术人员都清楚如何选择适当的载体、启动子、增强子和宿主细胞。

用重组 DNA 转化宿主细胞可用本领域技术人员熟知的常规技术进行。当宿主为原核生物如大肠杆菌时，能吸收 DNA 的感受态细胞可在指数生长期后收获，用 CaCl₂ 法处理，所用的步骤在本领域众所周知。另一种方法是使用 MgCl₂。如果需要，转化也可用电穿孔的方法进行。当宿主是真核生物，可选用如下的 DNA 转染方法：磷酸钙共沉淀法，常规机械方法如显微注射、电穿孔、脂质体包装等。

获得的转化子可以用常规方法培养，表达本发明的基因所编码的多肽。根据所用的宿主细胞，培养中所用的培养基可选自各种常规培养基。在适于宿主细胞生长的条件下进行培养。当宿主细胞生长到适当的细胞密度后，用合适的方法(如温度转换或化学诱导)诱导选择的启动子，将细胞再培养一段时间。

在上面的方法中的重组多肽可在细胞内、或在细胞膜上表达、或分泌到细胞外。如果需要，可利用其物理的、化学的和其它特性通过各种分离方法分离和纯化重组的蛋白。这些方法是本领域技术人员所熟知的。这些方法的例子包括但并不限于：常规的复性处理、用蛋白沉淀剂处理(盐析方法)、离心、渗透破菌、超处理、超离心、分子筛层析(凝胶过滤)、吸附层析、离子交换层析、高效液相层析(HPLC)和其

它各种液相层析技术及这些方法的结合。

重组的 Rhor 蛋白或多肽有多方面的用途。这些用途包括(但不限于): 直接做为药物治疗 Rhor 蛋白功能低下或丧失所致的疾病(如无精子, 或精子活力低下等疾病), 和用于筛选促进或对抗 Rhor 蛋白功能的抗体、多肽或其它配体。用表达的重组 Rhor 蛋白筛选多肽库可用于寻找有治疗价值的能抑制或刺激 Rhor 蛋白功能的多肽分子。

另一方面, 本发明还包括对 Rhor DNA 或是其片段编码的多肽具有特异性的多克隆抗体和单克隆抗体, 尤其是单克隆抗体。这里, “特异性”是指抗体能结合于 Rhor 基因产物或片段。较佳地, 指那些能与 Rhor 基因产物或片段结合但不识别和结合于其它非相关抗原分子的抗体。本发明中抗体包括那些能够结合并抑制 Rhor 蛋白的分子, 也包括那些并不影响 Rhor 蛋白功能的抗体。本发明还包括那些能与修饰或未经修饰形式的 Rhor 基因产物结合的抗体。

本发明不仅包括完整的单克隆或多克隆抗体, 而且还包括具有免疫活性的抗体片段, 如 Fab'或(Fab)₂ 片段; 抗体重链; 抗体轻链; 遗传工程改造的单链 Fv 分子; 或嵌合抗体, 如具有鼠抗体结合特异性但仍保留来自人的抗体部分的抗体。

本发明的抗体可以通过本领域内技术人员已知的各种技术进行制备。例如, 纯化的 Rhor 基因产物或者其具有抗原性的片段, 可被施用于动物以诱导多克隆抗体的产生。与之相似的, 表达 Rhor 蛋白或其具有抗原性的片段的细胞可用来免疫动物来生产抗体。本发明的抗体也可以是单克隆抗体。此类单克隆抗体可以利用杂交瘤技术来制备。本发明的抗体包括能阻断 Rhor 蛋白功能的抗体以及不影响 Rhor 蛋白功能的抗体。本发明的各类抗体可以利用 Rhor 基因产物的片段或功能区, 通过常规免疫技术获得。这些片段或功能区可以利用重组方法制备或利用多肽合成仪合成。与 Rhor 基因产物的未修饰形式结合的抗体可以用原核细胞(例如 *E. Coli*)中生产的基因产物来免疫动物而产生; 与翻译后修饰形式结合的抗体(如糖基化或磷酸化的蛋白或多肽), 可以用真核细胞(例如酵母或昆虫细胞)中产生的基因产物来免疫动物而获得。

抗 Rhor 蛋白的抗体可用于免疫组织化学技术中, 检测活检标本中的 Rhor 蛋白。此外, 与 Rhor 蛋白结合的单克隆抗体也可用放射性同位素标记, 注入体内可跟踪其位置和分布。

多克隆抗体的生产可用 Rhor 蛋白或多肽免疫动物, 如家兔, 小鼠, 大鼠等。多种佐剂可用于增强免疫反应, 包括但不限于弗氏佐剂等。

利用本发明蛋白, 通过各种常规筛选方法, 可筛选出与 Rhor 蛋白发生相互作用的物质, 如受体、抑制剂、激动剂或拮抗剂等。

本发明蛋白及其抗体、抑制剂、激动剂、拮抗剂或受体等, 当在治疗上进行施用(给药)时, 可提供不同的效果。通常, 可将这些物质配制于无毒的、惰性的和药

学上可接受的水性载体介质中，其中 pH 通常约为 5-8，较佳地 pH 约为 6-8，尽管 pH 值可随被配制物质的性质以及待治疗的病症而有所变化。配制好的药物组合物可以通过常规途径进行给药，其中包括(但并不限于)：口服、肌内、腹膜内、静脉内、皮下、皮内、或局部给药。

本发明的多肽可直接用于疾病治疗，例如，用于秃发的治疗。在使用本发明 Rhor 蛋白时，还可同时使用其他用于同一病症的治疗剂。

本发明还提供了一种药物组合物，它含有安全有效量的本发明 Rhor 多肽或其激动剂、拮抗剂以及药学上可接受的载体或赋形剂。这类载体包括(但并不限于)：盐水、缓冲液、葡萄糖、水、甘油、乙醇、及其组合。药物制剂应与给药方式相匹配。本发明的药物组合物可以被制成针剂形式，例如用生理盐水或含有葡萄糖和其他辅剂的水溶液通过常规方法进行制备。诸如片剂和胶囊之类的药物组合物，可通过常规方法进行制备。药物组合物如针剂、溶液、片剂和胶囊宜在无菌条件下制造。活性成分的给药量是治疗有效量，例如每天约 1 微克/千克体重-约 5 毫克/千克体重。此外，本发明的多肽还可与其他治疗剂一起使用。

使用药物组合物时，是将安全有效量的 Rhor 蛋白或其拮抗剂、激动剂施用于哺乳动物，其中该安全有效量通常至少约 10 微克/千克体重，而且在大多数情况下不超过约 5 毫克/千克体重，较佳地该剂量是约 10 微克/千克体重-约 1 毫克/千克体重。当然，具体剂量还应考虑给药途径、病人健康状况等因素，这些都是熟练医师技能范围内的。

Rhor 蛋白的多聚核苷酸也可用于多种治疗目的。基因治疗技术可用于治疗由于 Rhor 蛋白的无表达或异常/无活性的 Rhor 蛋白的表达所致的秃发。来源于病毒的表达载体如逆转录病毒、腺病毒、腺病毒相关病毒、单纯疱疹病毒、细小病毒等可用于将 Rhor 基因转移至细胞内。另外重组 Rhor 基因可包装到脂质体中，然后再转移至细胞内。

多聚核苷酸导入组织或细胞内的方法包括：将多聚核苷酸直接注入到体内组织中；或在体外通过载体(如病毒、噬菌体或质粒等)先将多聚核苷酸导入细胞中，再将细胞移植到体内等。

能与 Rhor 蛋白结合的多肽分子可通过筛选由各种可能组合的氨基酸结合于固相物组成的随机多肽库而获得。筛选时，必须对 Rhor 蛋白分子进行标记。

本发明还涉及定量和定位检测 Rhor 蛋白水平的诊断试验方法。这些试验是本领域所熟知的，且包括 FISH 测定和放射免疫测定。试验中所检测的 Rhor 蛋白水平，可以用作解释 Rhor 蛋白在各种疾病中的重要性和用于诊断 Rhor 蛋白起作用的疾病(如秃发)。

一种检测检测样品中是否存在 Rhor 蛋白的方法是利用 Rhor 蛋白的特异性抗体进行检测，它包括：将样品与 Rhor 蛋白特异性抗体接触；观察是否形成抗体复合

物，形成了抗体复合物就表示样品中存在Rhor蛋白。

Rhor蛋白的多聚核苷酸可用于Rhor蛋白相关疾病的诊断和治疗。在诊断方面，Rhor蛋白的多聚核苷酸可用于检测Rhor蛋白的表达与否或在疾病状态下Rhor蛋白的异常表达。如Rhor DNA序列可用于对活检标本的杂交以判断Rhor蛋白的表达异常。杂交技术包括Southern印迹法，Northern印迹法、原位杂交等。这些技术方法都是公开的成熟技术，相关的试剂盒都可从商业途径得到。本发明的多核苷酸的一部分或全部可作为探针固定在微阵列(microarray)或DNA芯片(又称为“基因芯片”)上，用于分析组织中基因的差异表达分析和基因诊断。用Rhor蛋白特异的引物进行RNA-聚合酶链反应(RT-PCR)体外扩增也可检测Rhor蛋白的转录产物。

检测Rhor基因的突变也可用于诊断Rhor蛋白相关的疾病。Rhor蛋白突变的形式包括与正常野生型Rhor DNA序列相比的点突变、易位、缺失、重组和其它任何异常等。可用已有的技术如Southern印迹法、DNA序列分析、PCR和原位杂交检测突变。另外，突变有可能影响蛋白的表达，因此用Northern印迹法、Western印迹法可间接判断基因有无突变。本发明的试验已表明，Rhor的突变与秃发直接相关。

在本发明的一个实例中，根据无毛、稀毛和正常小鼠的表型和遗传分析等手段，确定了一个从未被报导过的新基因，并分离了相应的多核苷酸序列，它编码具有SEQ ID NO: 2所示氨基酸序列的多肽。本发明的多核苷酸的序列如SEQ ID NO: 1所示，它包含的多核苷酸序列全长为2484个碱基，其开放读框位于1-2481位，编码全长为827个氨基酸的Rhor蛋白(SEQ ID NO: 2)。研究表明，该基因的突变可能阻断了信号的传导，影响了毛囊细胞的发育，从而导致了小鼠的先天性无毛。由于该蛋白的编码区存在高度的保守性，因此人类基因组内该基因的部分缺失极有可能导致人类的先天性秃发、稀发。Rhor为治疗秃发等疾病提供新的治疗途径，因而具有巨大的应用前景。

下面结合具体实施例，进一步阐述本发明。应理解，这些实施例仅用于说明本发明而并不用于限制本发明的范围。下列实施例中未注明具体条件的实验方法，通常按照常规条件如Sambrook等人，分子克隆：实验室手册(New York: Cold Spring Harbor Laboratory Press, 1989)中所述的条件，或按照制造厂商所建议的条件。

实施例 1

无毛相关基因的定位

正常小鼠Balb/c在培养中产生了先天性无毛的突变，杂合子表型为稀毛，纯合子表型为无毛。无毛小鼠与正常小鼠的杂交后代无一例外的表现为稀毛。稀毛

小鼠自交后代出现了性状分离且符合典型的孟德尔单基因遗传规律，其比例符合计算值即：正常：稀毛：无毛=1：2：1。此外，根据其系谱确定该种突变是单基因的不完全显性。

为了检测该种突变，抽提 2000 只小鼠的基因组 DNA 进行 genotyping 所得的数据用计算机进行分析，所用软件为 Linkage，将突变定位于微卫星标记 D11mit103 和 D11mit338 之间，它们的物理距离为 1700kb，随后构建了该区域精细的物理图谱，进一步将突变定位到了 800kb 的区域，然后对该区域内的基因进行基于直接测序方法的全基因扫描。从而将突变位点定位到一个较小的区域内。

实施例 2

基因组 DNA 缺失的鉴定：

用分子克隆中所述的标准方法从无毛、稀毛和正常 Balb/c 小鼠组织中抽提基因组 DNA，使用以下合成引物

正向引物：5' gcaggctagcgtgttaaagg 3' (SEQ ID NO:3)

反向引物：5' aaaacgggggtcatagcagc 3' (SEQ ID NO:4)

以基因组 DNA 为模板使用 ABI 公司的 TaqGold 高温聚合酶进行 PCR 扩增，PCR 热循环条件是：

高温聚合酶热启动，模板变性：95℃ 10 分钟

第一轮热循环

95℃ 30 秒

68℃ 45 秒

每一次循环降低一度

72℃ 2 分 30 秒

循环 12 次

第二轮热循环

95℃ 30 秒

56℃ 45 秒

72℃ 2 分 30 秒

循环 35 次

72℃ 10 分钟

产物进行 1.0% 琼脂糖凝胶电泳，溴化乙锭染色结果如图 1 所示。杂合稀毛小鼠的基因组扩增产物为两条特异性条带其分子量与无毛和正常小鼠的扩增产物分别吻合。

实施例 3

Rhor 的基因组测序

本实施例进一步通过基因组测序的方法，验证 Rhor 与无毛的相关性。方法如下：

用 Omega 公司的 Cycle-Pure Kit 纯化上述无毛与正常小鼠基因组 PCR 产物，其中使用以下合成引物

正向引物： 5' gcacatctgaggggaaggaag 3' (SEQ ID NO: 5)

反向引物： 5' cttccggtcaaagtgaaagt 3' (SEQ ID NO: 6)

用 ABI 公司的 BigDye terminal 试剂及 3100 全自动测序仪，测序反应条件是：

模板变性 98 度 2 分

96 度 20 秒

51 度 20 秒

60 度 4 分

循环 25 次

测序结果如下所示：

```

AGCCTACCTG AAGAGTGTCA GCCTACAGGA GCGCCGGGGA CGATGGCAGG AGGGCGCAGA 60
GAAGCGCCCC GGCTTCGCGC GCCAGGCCTC CCTGTCCCAG AGCATCCGCA AGAGCACAGC 120
CCAGTGGT TTT GGGGTCAGCG GCGACTGGGA GGGCAAGCGA CAAAAGTGGC ATCGTCGCAG 180
CCTGCACCAC TGCAGCGTGC ACTATGGCCG CCTCAAGGCC TCGTGCCAGA GAGAACTGGA 240
GCTGCCCAGC CAGGAGGTGC CATCCTTCCA GGGCACTGAG TCTCCAAAAC CGTGCAAGAT 300
GCCAAGATT GTGGATCCAC TGGCTCGGGG TAGGGCCTTC CGCCATCCAG ATGAGGTGGA 360
CCGGCCTCAC GCTGCCCACC CACCTCTGAC TCCAGGGGTC CTGTCTCTCA CATCCTTCAC 420
CATGTCCGCT CTGGCTACTC CCATCTGCCC CGCCGCAAGA GGATATCTGT TGCCCATATG 480
AGCTTTCAGG CAGCCGCCGC CCTCCTCAAG GGGCGTTCCG TGCTAGATGC GACTGGGCAG 540
CGGTGCCGGC ATGTCAAACG CAGCTTCGCT TACCCAGCT TCCTGGAGGA GGATGCTGTC 600
(SEQ ID NO: 7)
  
```

将正常小鼠的序列与无毛小鼠的序列比较可发现，无毛小鼠的基因组发生了 230 碱基的缺失（下划线部分）。

实施例 4

Rhor 基因完整开放阅读框架的获得

按照 Qiagen 公司的 Total Rna 纯化试剂盒提供的方法抽提小鼠皮肤总 RNA，1% 琼脂糖电泳鉴定纯度后用 Promega 公司的 Reverse Transcript System 将

RNA 反转录为 cDNA. 后用以下引物:

正向引物: 5' ACTCTGCTCTCAGCCGCTT 3' (SEQ ID NO: 8)

反向引物: 5' CCAGACACATGCTGGAGCTA 3' (SEQ ID NO: 9)

使用 Takara 公司的 LA 高温聚合酶

PCR 热循环条件:

95 度 30 秒

54 度 45 秒

72 度 3 分

循环 45 次

72 度 10 分钟

0.8% 琼脂糖电泳分离, 割胶纯化, 测序鉴定, 得到 Rhor 的完整 ORF 全长序列。

Rhor 基因的完整开放阅读框架如 SEQ ID NO:1 所示。其编码蛋白的氨基酸序列如 SEQ ID NO:2 所示。

实施例 5

Rhor 基因 cDNA 缺失的鉴定方法

使用上述方法获得的无毛及正常小鼠 cDNA 为模板, 使用 ABI 公司的 GoldTaq 高温聚合酶。引物如下:

正向引物: 5' gcacatctgaggggaaggaag 3' (SEQ ID NO: 10)

反向引物: 5' cttccggtcaaagtgaaagt 3' (SEQ ID NO: 11)

PCR 热循环条件

模板热变性, 高温聚合酶热启动 95 度 10 分

第一轮热循环

95 度 30 秒

68 度 45 秒

每一循环降低 1 度

72 度 90 秒

循环 14 次

第二轮热循环

95 度 30 秒

54 度 45 秒

72 度 90 秒

循环 40 次

72 度 10 分

0.8%琼脂糖电泳鉴定后用 Omega 公司的 Cycle-Pure Kit 纯化产物, 用 ABI 公司的 BigDye terminal 仍用上述引物测序, 测序反应:

模板变性 98 度 2 分

96 度 20 秒

51 度 20 秒

60 度 4 分

循环 25 次

结果如下

```
GTGTCAGCCT ACAGGAGCCC CGGGGACGAT GGCAGGAGGG CGCAGAGAAG CGCCCCGGCT 60
TCCGCCGCCA GGCTCCCTG TCCAGAGCA TCCGCAAGAG CACAGCCCAG TGGTTTGGGG 120
TCAGCGGCGA CTGGGAGGGC AAGCGACAAA ACTGGCATCG TCGCAGCCTG CACCACTGCA 180
GCGTGACTA TGGCCGCTC AAGGCCTCGT GCCAGAGAGA ACTGGAGCTG CCCAGCCAGG 240
AGGTGCCATC CTTCCAGGGC ACTGAGTCTC CAAAACCGTG CAAGATGCCC AAGATTGTGG 300
ATCCACTGGC TCGGGGTAGG GCCTTCCGCC ATCCAGATGA GGTGGACCGG CCTCACGCTG 360
CCCACCCACC TCTGACTCCA GGGGTCCTGT CTCTCACATC CTTACCATG TCCGCTCTGG 420
CTACTCCCAT CTGCCCCGCC GCAAGAGGAT ATCTGTTGCC CATATGAGCT TTCAGGCAGC 480
CGCCGCCCTC CTCAAGGGGC GTTCCGTGCT AGATGCGACT GGGCAGCGGT GCCGGCATGT 540
CAAACGCAGC TTCGCTTACC CCAGCTTCCT GGAGGAGGAT GCTGTCGATG GAGCTGACAC 600
(SEQ ID NO: 12)
```

其中, 下划线部分为无毛小鼠与正常小鼠比较后发生缺失的序列。其中 cDNA 序列中的缺失部分与基因组序列中的缺失部分相同。

实施例 6

Rhor 的结构分析

通过结构分析, 发现 Rhor 基因编码的蛋白序列包含了二个非常保守的结构域:

(1) Rhor 蛋白的氨基酸 619—759 位形成一个 Rhomboid 结构域。这表明, Rhor 蛋白属于 Rhomboid 蛋白家族成员, 该家族成员在细菌和真核生物中都有报道, 大部分是完整的膜蛋白, 它们预测跨膜区存在三个高度保守的组氨酸。

(2) Rhor 蛋白的氨基酸 610—804 位形成一个 competence 结构域。这表明, 它属于 competence 蛋白家族成员, 该家族成员是完整的膜蛋白, 它们具有 6 个跨膜的螺旋结构, 这些蛋白行使跨膜运输核酸的功能。一些家族成员曾被报导在细菌吸收细胞外 DNA 功能中起了关键的作用。该家族成员有一个氨基末端的跨膜区, 该区内两个组氨酸形成一个高度保守的结构单元 (motif) 作为金属离子结合位点。

上述结构分析提示, Rhor 蛋白的突变可能使其失去了跨核膜运输信使 RNA (mRNA) 的功能, 从而阻断了毛囊细胞的发育。

实施例 7

Rhor 的其他功能

除了无毛之外，还发现无毛小鼠普遍较正常小鼠要消瘦，在解剖的过程中发现其脂肪储备几乎不可见。这表明，认为 Rhor 基因也参与了脂肪代谢通路的调节，它的突变导致了脂代谢的紊乱。

此外，无毛小鼠在生长过程中出现了淋巴瘤，Rhor 基因的突变会导致淋巴细胞的生长失去正常调控。这暗示 Rhor 有可能处于多个代谢途径的调控点。

实施例 8

Rhor 在原核细胞中的表达

合成以下引物：

正向引物： 5'CGGATCCATGGCCTCAGCTGACAAGAATGGCAGCAACCTCCCA 3'(SEQ ID NO:13) (引入 *Bam*HI 切点 GGATCC, 第一个 C 为保护碱基)

反向引物： 5'ATAAAGCTTGCTCGATCTGGTCCACGATGTGATT 3'(SEQ ID NO: 14)(引入 *Hind*III 切点 AAGCTT, 起始 ATA 为保护碱基)

以小鼠 cDNA 为模板, PCR 扩增编码 Rhor 蛋白的 DNA。

利用 *Bam*HI 和 *Hind* III 位点, 将 PCR 扩增产物克隆到相同酶切的质粒 pET32a (Novagen 公司) 中, 然后转化宿主 *E. Coli:BL21 (DE3)*。转化的宿主细胞在 IPTG 的诱导下表达。

表达产物在 SDS-PAGE 胶中分子量大约为 90Kda。

实施例 9

抗体的制备

用实施例 8 中获得蛋白用作抗原, 在兔中制备多克隆抗体。具体程序如下：第一次免疫纯种新西兰兔, 蛋白用量为 240ug/只, 用弗氏完全佐剂。4 周后, 第二次免疫, 蛋白用量为 140ug/只, 用不完全佐剂。之后 2 周, 第三次免疫, 蛋白用量为 120ug/只, 用不完全佐剂。

在本发明提及的所有文献都在本申请中引用作为参考, 就如同每一篇文献被单独引用作为参考那样。此外应理解, 在阅读了本发明的上述讲授内容之后, 本领域技术人员可以对本发明作各种改动或修改, 这些等价形式同样落于本申请所附权利要求书所限定的范围。

权 利 要 求

- 1.一种分离的 Rhor 多肽,其特征在于,它含有 SEQ ID NO: 2 氨基酸序列的多肽。
- 2.如权利要求 1 所述的多肽,其特征在于,该多肽是具有 SEQ ID NO: 2 氨基酸序列的多肽。
- 3.一种分离的多核苷酸,其特征在于,它包含一核苷酸序列,该核苷酸序列与选自下组:
 - (a)编码含有 SEQ ID NO:2 氨基酸序列 的多肽的多核苷酸;
 - (b)与多核苷酸(a)互补的多核苷酸。
- 4.如权利要求 3 所述的多核苷酸,其特征在于,该多核苷酸编码具有 SEQ ID NO: 2 所示氨基酸序列的多肽。
- 5.如权利要求 3 所述的多核苷酸,其特征在于,该多核苷酸具有 SEQ ID NO: 1 中 1-2484 位的序列。
- 6.一种载体,其特征在于,它含有权利要求 3 所述的多核苷酸。
- 7.一种遗传工程化的宿主细胞,其特征在于,它含有权利要求 6 所述的载体。
- 8.一种 Rhor 蛋白的制备方法,其特征在于,该方法包含:
 - (a)在表达 Rhor 蛋白的条件下,培养权利要求 7 所述的宿主细胞;
 - (b)从培养物中分离出 Rhor 蛋白。
- 9.一种检测秃发易感性的试剂盒,其特征在于,它包括特异性扩增 Rhor 基因或转录本的引物。
- 10.一种组合物,其特征在于,它含有安全有效量的权利要求 1 所述的多肽以及药学上可接受的载体。

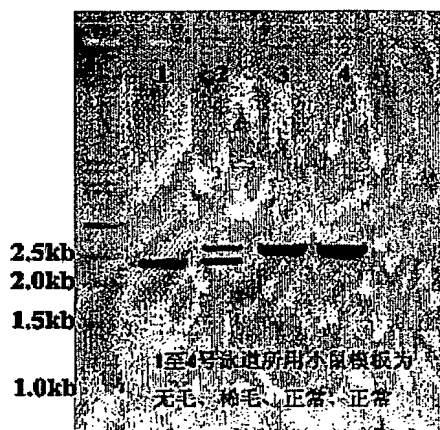


图 1

序列表

<110> 中国科学院上海生命科学研究院

<120> 利用人和鼠 Rhor 基因及其编码产物诊断和治疗秃发的方法

<130> 026816 pc

<150> CN 02145353.9

<151> 2002-11-13

<160> 14

<170> PatentIn version 3.1

<210> 1

<211> 2484

<212> DNA

<213> 小鼠 (Mus musculus)

<220>

<221> CDS

<222> (1)..(2481)

<223>

<400> 1

atg gcc tca gct gac aag aat ggc agc aac ctc cca tct gtg tct ggt	48
Met Ala Ser Ala Asp Lys Asn Gly Ser Asn Leu Pro Ser Val Ser Gly	
1 5 10 15	
agc cgc ctg cag agc cgg aag cca ccc aac ctc tcc atc acc atc ccg	96
Ser Arg Leu Gln Ser Arg Lys Pro Pro Asn Leu Ser Ile Thr Ile Pro	
20 25 30	
cca cca gag agc cag gcc ccc ggc gag cag gat agc atg ctt cct gag	144
Pro Pro Glu Ser Gln Ala Pro Gly Glu Gln Asp Ser Met Leu Pro Glu	
35 40 45	
agg cgc aag aac cca gcc tac ctg aag agt gtc agc cta cag gag ccc	192
Arg Arg Lys Asn Pro Ala Tyr Leu Lys Ser Val Ser Leu Gln Glu Pro	
50 55 60	
cgg gga cga tgg cag gag ggc gca gag aag cgc ccc ggc ttc cgc cgc	240
Arg Gly Arg Trp Gln Glu Gly Ala Glu Lys Arg Pro Gly Phe Arg Arg	
65 70 75 80	
cag gcc tcc ctg tcc cag agc atc cgc aag agc aca gcc cag tgg ttt	288
Gln Ala Ser Leu Ser Gln Ser Ile Arg Lys Ser Thr Ala Gln Trp Phe	
85 90 95	
ggg gtc agc ggc gac tgg gag ggc aag cga caa aac tgg cat cgt cgc	336
Gly Val Ser Gly Asp Trp Glu Gly Lys Arg Gln Asn Trp His Arg Arg	
100 105 110	
agc ctg cac cac tgc agc gtg cac tat ggc cgc ctc aag gcc tcg tgc	384
Ser Leu His His Cys Ser Val His Tyr Gly Arg Leu Lys Ala Ser Cys	
115 120 125	
cag aga gaa ctg gag ctg ccc agc cag gag gtg cca tcc ttc cag ggc	432
Gln Arg Glu Leu Glu Leu Pro Ser Gln Glu Val Pro Ser Phe Gln Gly	
130 135 140	
act gag tct cca aaa ccg tgc aag atg ccc aag att gtg gat cca ctg	480
Thr Glu Ser Pro Lys Pro Cys Lys Met Pro Lys Ile Val Asp Pro Leu	
145 150 155 160	
gct cgg ggt agg gcc ttc cgc cat cca gat gag gtg gac cgg cct cac	528
Ala Arg Gly Arg Ala Phe Arg His Pro Asp Glu Val Asp Arg Pro His	
165 170 175	
gct gcc cac cca cct ctg act cca ggg gtc ctg tct ctc aca tcc ttc	576

2/7

Lys	Trp	Gln	Asn	Asp	Thr	Gly	Pro	Ser	Asp	Lys	Ser	Asp	Leu	Ser	Gln		
			500					505					510				
aag	cag	cca	tcg	gcg	gtt	gtg	tgc	cac	caa	gac	ccc	agg	acc	tgt	gaa	1584	
Lys	Gln	Pro	Ser	Ala	Val	Val	Cys	His	Gln	Asp	Pro	Arg	Thr	Cys	Glu		
		515					520					525					
gag	cct	gcc	tcc	agt	ggg	gcc	cac	atc	tgg	cct	gat	gac	att	acc	aag	1632	
Glu	Pro	Ala	Ser	Ser	Gly	Ala	His	Ile	Trp	Pro	Asp	Asp	Ile	Thr	Lys		
		530				535					540						
tgg	cgc	atc	tgc	aca	gag	cag	gct	cag	agc	aac	cac	acg	ggc	ttg	ttg	1680	
Trp	Pro	Ile	Cys	Thr	Glu	Gln	Ala	Gln	Ser	Asn	His	Thr	Gly	Leu	Leu		
545					550					555					560		
cac	ata	gac	tgt	aag	atc	aaa	ggc	cgc	ccc	tgc	tgc	atc	ggc	acc	aag	1728	
His	Ile	Asp	Cys	Lys	Ile	Lys	Gly	Arg	Pro	Cys	Cys	Ile	Gly	Thr	Lys		
			565					570					575				
ggc	agc	tgc	gag	atc	acc	act	cgg	gag	tac	tgt	gag	ttc	atg	cat	ggc	1776	
Gly	Ser	Cys	Glu	Ile	Thr	Thr	Arg	Glu	Tyr	Cys	Glu	Phe	Met	His	Gly		
		580						585				590					
tat	ttc	cat	gaa	gac	gcg	acg	ctg	tgt	tcc	cag	gtg	cac	tgt	tta	gac	1824	
Tyr	Phe	His	Glu	Asp	Ala	Thr	Leu	Cys	Ser	Gln	Val	His	Cys	Leu	Asp		
		595					600					605					
aag	gtg	tgt	ggg	ctc	ctg	cct	ttc	ctc	aac	cct	gag	gtc	cct	gac	cag	1872	
Lys	Val	Cys	Gly	Leu	Leu	Pro	Phe	Leu	Asn	Pro	Glu	Val	Pro	Asp	Gln		
	610					615					620						
ttc	tac	cgg	atc	tgg	ctg	tct	tta	ttc	ctg	cat	gct	ggc	ata	gtg	cac	1920	
Phe	Tyr	Arg	Ile	Trp	Leu	Ser	Leu	Phe	Leu	His	Ala	Gly	Ile	Val	His		
625					630					635					640		
tgc	ctt	gtg	tct	gtg	gtc	ttc	caa	atg	acc	atc	ctg	agg	gac	cta	gag	1968	
Cys	Leu	Val	Ser	Val	Val	Phe	Gln	Met	Thr	Ile	Leu	Arg	Asp	Leu	Glu		
			645					650				655					
aag	ctg	gcc	ggc	tgg	cac	cgc	atc	tcc	atc	atc	ttc	atc	ctt	agt	ggc	2016	
Lys	Leu	Ala	Gly	Trp	His	Arg	Ile	Ser	Ile	Ile	Phe	Ile	Leu	Ser	Gly		
		660					665					670					
att	aca	ggc	aac	ctg	gcc	agc	gcc	atc	ttc	ctc	ccc	tac	cgg	gca	gag	2064	
Ile	Thr	Gly	Asn	Leu	Ala	Ser	Ala	Ile	Phe	Leu	Pro	Tyr	Arg	Ala	Glu		
		675					680					685					
gtg	ggc	cca	gcc	ggg	tcg	cag	ttc	ggc	ctc	ctc	ggc	tgc	ctc	ttc	gtg	2112	
Val	Gly	Pro	Ala	Gly	Ser	Gln	Phe	Gly	Leu	Leu	Ala	Cys	Leu	Phe	Val		
	690					695					700						
gag	ctg	ttc	cag	agc	tgg	cag	ctg	ttg	gag	cgg	ccg	tgg	aag	gcc	ttc	2160	
Glu	Leu	Phe	Gln	Ser	Trp	Gln	Leu	Leu	Glu	Arg	Pro	Trp	Lys	Ala	Phe		
705					710					715					720		
ttc	aac	ctg	tcg	gcc	att	gtg	ctt	ttc	ctc	ttc	atc	tgt	ggc	ctc	ctg	2208	
Phe	Asn	Leu	Ser	Ala	Ile	Val	Leu	Phe	Leu	Phe	Ile	Cys	Gly	Leu	Leu		
			725					730					735				
ccc	tgg	ata	gac	aac	atc	gcc	cac	atc	ttc	ggg	ttc	ctc	agc	ggc	atg	2256	
Pro	Trp	Ile	Asp	Asn	Ile	Ala	His	Ile	Phe	Gly	Phe	Leu	Ser	Gly	Met		
		740					745					750					
ctt	ctg	gcc	ttc	gcc	ttc	ctg	cct	tac	att	acc	ttc	ggc	acc	agc	gac	2304	
Leu	Leu	Ala	Phe	Ala	Phe	Leu	Pro	Tyr	Ile	Thr	Phe	Gly	Thr	Ser	Asp		
		755					760					765					
aag	tac	cgc	aag	cga	gcc	ctc	atc	ctc	gtg	tcg	ctg	ctg	gtc	ttt	gct	2352	
Lys	Tyr	Arg	Lys	Arg	Ala	Leu	Ile	Leu	Val	Ser	Leu	Leu	Val	Phe	Ala		
	770					775					780						
ggg	ctc	ttt	gct	tcc	ctg	gtg	ctg	tgg	ctg	tac	atc	tac	ccc	atc	aac	2400	
Gly	Leu	Phe	Ala	Ser	Leu	Val	Leu	Trp	Leu	Tyr	Ile	Tyr	Pro	Ile	Asn		
785					790					795					800		
tgg	ccc	tgg	atc	gag	tac	ctc	acc	tgc	ttt	ccc	ttc	acc	agc	cgc	ttc	2448	
Trp	Pro	Trp	Ile	Glu	Tyr	Leu	Thr	Cys	Phe	Pro	Phe	Thr	Ser	Arg	Phe		
			805					810					815				
tgt	gag	aag	tac	gag	cta	gac	cag	gtg	cta	cac	taa					2484	

Cys Glu Lys Tyr Glu Leu Asp Gln Val Leu His
820 825

<210> 2
<211> 827
<212> PRT
<213> 小鼠 (Mus musculus)

<400> 2

```

Met Ala Ser Ala Asp Lys Asn Gly Ser Asn Leu Pro Ser Val Ser Gly
1      5      10      15
Ser Arg Leu Gln Ser Arg Lys Pro Pro Asn Leu Ser Ile Thr Ile Pro
20      25      30
Pro Pro Glu Ser Gln Ala Pro Gly Glu Gln Asp Ser Met Leu Pro Glu
35      40      45
Arg Arg Lys Asn Pro Ala Tyr Leu Lys Ser Val Ser Leu Gln Glu Pro
50      55      60
Arg Gly Arg Trp Gln Glu Gly Ala Glu Lys Arg Pro Gly Phe Arg Arg
65      70      75      80
Gln Ala Ser Leu Ser Gln Ser Ile Arg Lys Ser Thr Ala Gln Trp Phe
85      90      95
Gly Val Ser Gly Asp Trp Glu Gly Lys Arg Gln Asn Trp His Arg Arg
100     105     110
Ser Leu His His Cys Ser Val His Tyr Gly Arg Leu Lys Ala Ser Cys
115     120     125
Gln Arg Glu Leu Glu Leu Pro Ser Gln Glu Val Pro Ser Phe Gln Gly
130     135     140
Thr Glu Ser Pro Lys Pro Cys Lys Met Pro Lys Ile Val Asp Pro Leu
145     150     155     160
Ala Arg Gly Arg Ala Phe Arg His Pro Asp Glu Val Asp Arg Pro His
165     170     175
Ala Ala His Pro Pro Leu Thr Pro Gly Val Leu Ser Leu Thr Ser Phe
180     185     190
Thr Ser Val Arg Ser Gly Tyr Ser His Leu Pro Arg Arg Lys Arg Ile
195     200     205
Ser Val Ala His Met Ser Phe Gln Ala Ala Ala Leu Leu Lys Gly
210     215     220
Arg Ser Val Leu Asp Ala Thr Gly Gln Arg Cys Arg His Val Lys Arg
225     230     235     240
Ser Phe Ala Tyr Pro Ser Phe Leu Glu Glu Asp Ala Val Asp Gly Ala
245     250     255
Asp Thr Phe Asp Ser Ser Phe Phe Ser Lys Glu Glu Met Ser Ser Met
260     265     270
Pro Asp Asp Val Phe Glu Ser Pro Pro Leu Ser Ala Ser Tyr Phe Arg
275     280     285
Gly Val Pro His Ser Ala Ser Pro Val Ser Pro Asp Gly Val His Ile
290     295     300
Pro Leu Lys Glu Tyr Ser Gly Gly Arg Ala Leu Gly Pro Gly Thr Gln
305     310     315     320
Arg Gly Lys Arg Ile Ala Ser Lys Val Lys His Phe Ala Phe Asp Arg
325     330     335
Lys Lys Arg His Tyr Gly Leu Gly Val Gly Asn Trp Leu Asn Arg
340     345     350
Ser Tyr Arg Arg Ser Ile Ser Ser Thr Val Gln Arg Gln Leu Glu Ser
355     360     365
Phe Asp Ser His Arg Pro Tyr Phe Thr Tyr Trp Leu Thr Phe Val His
370     375     380
Ile Ile Ile Thr Leu Leu Val Ile Cys Thr Tyr Gly Ile Ala Pro Val
385     390     395     400

```

Gly	Phe	Ala	Gln	His	Val	Thr	Thr	Gln	Leu	Val	Leu	Lys	Asn	Arg	Gly
				405					410					415	
Val	Tyr	Glu	Ser	Val	Lys	Tyr	Ile	Gln	Gln	Glu	Asn	Phe	Trp	Ile	Gly
			420					425					430		
Pro	Ser	Ser	Ile	Asp	Leu	Ile	His	Leu	Gly	Ala	Lys	Phe	Ser	Pro	Cys
		435					440					445			
Ile	Arg	Lys	Asp	Gln	Gln	Ile	Glu	Gln	Leu	Val	Arg	Arg	Glu	Arg	Asp
	450					455					460				
Ile	Glu	Arg	Thr	Ser	Gly	Cys	Cys	Val	Gln	Asn	Asp	Arg	Ser	Gly	Cys
465				470						475					480
Ile	Gln	Thr	Leu	Lys	Lys	Asp	Cys	Ser	Glu	Thr	Leu	Ala	Thr	Phe	Val
			485						490					495	
Lys	Trp	Gln	Asn	Asp	Thr	Gly	Pro	Ser	Asp	Lys	Ser	Asp	Leu	Ser	Gln
			500					505					510		
Lys	Gln	Pro	Ser	Ala	Val	Val	Cys	His	Gln	Asp	Pro	Arg	Thr	Cys	Glu
		515					520					525			
Glu	Pro	Ala	Ser	Ser	Gly	Ala	His	Ile	Trp	Pro	Asp	Asp	Ile	Thr	Lys
	530					535					540				
Trp	Pro	Ile	Cys	Thr	Glu	Gln	Ala	Gln	Ser	Asn	His	Thr	Gly	Leu	Leu
545				550						555					560
His	Ile	Asp	Cys	Lys	Ile	Lys	Gly	Arg	Pro	Cys	Cys	Ile	Gly	Thr	Lys
			565					570						575	
Gly	Ser	Cys	Glu	Ile	Thr	Thr	Arg	Glu	Tyr	Cys	Glu	Phe	Met	His	Gly
			580					585					590		
Tyr	Phe	His	Glu	Asp	Ala	Thr	Leu	Cys	Ser	Gln	Val	His	Cys	Leu	Asp
		595					600					605			
Lys	Val	Cys	Gly	Leu	Leu	Pro	Phe	Leu	Asn	Pro	Glu	Val	Pro	Asp	Gln
	610					615					620				
Phe	Tyr	Arg	Ile	Trp	Leu	Ser	Leu	Phe	Leu	His	Ala	Gly	Ile	Val	His
625				630						635					640
Cys	Leu	Val	Ser	Val	Val	Phe	Gln	Met	Thr	Ile	Leu	Arg	Asp	Leu	Glu
			645					650						655	
Lys	Leu	Ala	Gly	Trp	His	Arg	Ile	Ser	Ile	Ile	Phe	Ile	Leu	Ser	Gly
			660					665					670		
Ile	Thr	Gly	Asn	Leu	Ala	Ser	Ala	Ile	Phe	Leu	Pro	Tyr	Arg	Ala	Glu
		675					680					685			
Val	Gly	Pro	Ala	Gly	Ser	Gln	Phe	Gly	Leu	Leu	Ala	Cys	Leu	Phe	Val
	690					695					700				
Glu	Leu	Phe	Gln	Ser	Trp	Gln	Leu	Leu	Glu	Arg	Pro	Trp	Lys	Ala	Phe
705					710					715					720
Phe	Asn	Leu	Ser	Ala	Ile	Val	Leu	Phe	Leu	Phe	Ile	Cys	Gly	Leu	Leu
			725					730						735	
Pro	Trp	Ile	Asp	Asn	Ile	Ala	His	Ile	Phe	Gly	Phe	Leu	Ser	Gly	Met
		740						745					750		
Leu	Leu	Ala	Phe												

<210>	3
<211>	20
<212>	DNA
<213>	引物

<400> 3	
gcaggctagc gtgttaaagg	20
<210> 4	
<211> 19	
<212> DNA	
<213> 引物	
<400> 4	
aaaacggggt catagcagc	19
<210> 5	
<211> 20	
<212> DNA	
<213> 引物	
<400> 5	
gcacatctga gggaaggaag	20
<210> 6	
<211> 20	
<212> DNA	
<213> 引物	
<400> 6	
cttccggtca aatgcaaagt	20
<210> 7	
<211> 600	
<212> DNA	
<213> 小鼠(Mus musculus)	
<400> 7	
agcctacctg aagagtgtca gcctacagga gcccggggga cgatggcagg agggcgcgaga	60
gaagcgcccc ggcttccgcc gccaggcctc cctgtcccag agcatccgca agagcacagc	120
ccagtgggtt ggggtcagcg gcgactggga gggcaagcga caaaactggc atcgctcgag	180
cctgcaccac tgcagcgtgc actatggccg cctcaaggcc tcgtgccaga gagaactgga	240
gtgtcccagc caggaggtgc catccttcca gggcactgag tctccaaaac cgtgcaagat	300
gcccagatt gtggatccac tggctcgggg tagggccttc cgccatccag atgaggtgga	360
ccggcctcac gctgccacc cacctctgac tccaggggtc ctgtctctca catccttcac	420
catgtccgct ctggctactc ccatctgccc cgccgcaaga ggatatctgt tgcccataatg	480
agctttcagg cagccgccg cctcctcaag gggcggttcg tgctagatgc gactgggcag	540
cggtgccggc atgtcaaag cagcttcgct taccacagct tcctggagga ggatgctgtc	600
<210> 8	
<211> 19	
<212> DNA	
<213> 引物	
<400> 8	
actctgctct cagccgctt	19
<210> 9	
<211> 20	
<212> DNA	
<213> 引物	
<400> 9	
ccagacacat gctggagcta	20

<210> 10
 <211> 20
 <212> DNA
 <213> 引物

<400> 10
 gcacatctga gggaaggaag 20

<210> 11
 <211> 20
 <212> DNA
 <213> 引物

<400> 11
 cttccggtca aatgcaaagt 20

<210> 12
 <211> 600
 <212> DNA
 <213> 小鼠 (Mus musculus)

<400> 12
 gtgtcagcct acaggagccc cggggacgat ggcaggaggg cgcagagaag cgccccggct 60
 tccgccgcca ggcttcctg tcccagagca tccgcaagag cacagcccag tggtttgggg 120
 tcagcggcga ctgggagggc aagcgacaaa actggcatcg tcgcagcctg caccactgca 180
 gcgtgcacta tggccgcctc aaggcctcgt gccagagaga actggagctg cccagccagg 240
 aggtgccatc cttccagggc actgagtctc caaaaccgtg caagatgccc aagattgtgg 300
 atccactggc tcggggtagg gccttcgcc atccagatga ggtggaccgg cctcacgctg 360
 cccaccacc tctgactcca ggggtcctgt ctctcacatc cttcaccatg tccgctctgg 420
 ctactcccat ctgccccgcc gcaagaggat atctgttgcc catatgagct ttcaggcagc 480
 cgccgccctc ctcaaggggc gttccgtgct agatgcgact gggcagcggt gccggcatgt 540
 caaacgcagc ttcgcttacc ccagcttcct ggaggaggat gctgtcgtg gagctgacac 600

<210> 13
 <211> 43
 <212> DNA
 <213> 引物

<400> 13
 cggatccatg gcctcagctg acaagaatgg cagcaacctc cca 43

<210> 14
 <211> 34
 <212> DNA
 <213> 引物

<400> 14
 ataaagcttg ctcgatctgg tccacgatgt gatt 34

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN03/00956

A. CLASSIFICATION OF SUBJECT MATTER

IPC⁷ : C07K14/47, C12N15/12, A61K38/17, A61P17/14

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI, EPODOC, PAJ, CNPAT, CNKI, GENE BANK

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	Genebank Online, accessiong no.: AK042635, 19 September 2003	1-10
P, X	Genebank Online, accessiong no.: AK028872, 18 September 2003	1-10
A	WO, A2, 0070945 (KAROLINSKA INNOVATIONS AB) 30 November 2000, See claims and figures	1-10
A	WO, A1, 9963083 (TAISHO PHARM CO LTD) 9 December 1999, See the whole document.	1-10

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim (S) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
22 March 2004 (22.03.04)

Date of mailing of the international search report
08 · APR 2004 (08 · 04 · 2004)

Name and mailing address of the ISA/CN
6 Xitucheng Rd., Jimen Bridge, Haidian District,
100088 Beijing, China
Facsimile No. 86-10-62019451

Authorized officer

CHANG, mao

Telephone No. 86-10-62085298



PCT/CN03/00956

Patent document cited in search report	Publication date	Patent family members	Publication date
WO, A2, 0070945	2000-11-30	AU, A, 200053941	2000-12-12
WO, A1, 9963083	1999-12-09	JP, A, 11332571 AU, A, 3955099	1999-12-07 1999-12-20

国际检索报告

PCT/CN03/00956

A. 主题的分类

IPC⁷: C07K14/47, C12N15/12, A61K38/17, A61P17/14

按照国际专利分类表(IPC)或者同时按照国家分类和 IPC 两种分类

B. 检索领域

检索的最低限度文献(标明分类体系和分类号)

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称和, 如果实际可行的, 使用的检索词)

WPI, EPODOC, PAJ, CNPAT, CNKI, Genbank

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求编号
P, X	Genebank Online, accession no: AK042365, 2003 年 9 月 19 日	1-10
P, X	Genebank Online, accession no.:AK028872, 2003 年 9 月 18 日	1-10
A	WO, A2, 0070945 (KAROLINSKA INNOVATIONS AB) 2000 年 11 月 30 日, 见权利要求和附图	1-10
A	WO, A1, 9963083 (大正制药株式会社) 1999 年 12 月 9 日, 见全文	1-10

☐ 其余文件在 C 栏的续页中列出。

☒ 见同族专利附件。

* 引用文件的专用类型:

“A” 明确叙述了被认为不是特别相关的一般现有技术的文件

“E” 在国际申请日的当天或之后公布的在先的申请或专利

“L” 可能引起对优先权要求的怀疑的文件, 为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布的在后文件, 它与申请不相抵触, 但是引用它是为了理解构成发明基础的理论或原理

“X” 特别相关的文件, 仅仅考虑该文件, 权利要求所记载的发明就不能认为是新颖的或不能认为是有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 权利要求记载的发明不具有创造性

“&” 同族专利成员的文件

国际检索实际完成的日期

22.3 月 2004 (22.03.04)

国际检索报告邮寄日期

08 · 4月 2004 (08 · 04 · 2004)

国际检索单位名称和邮寄地址

ISA/CN

中国北京市海淀区西土城路 6 号(100088)

传真号: 86-10-62019451

受权官员

常 矛



电话号码: 86-10-62085298

国际检索报告
关于同族专利成员的情报

国际申请号
PCT/CN03/00956

检索报告中引用的 专利文件	公布日期	同族专利成员	公布日期
WO, A2, 0070945	2000-11-30	AU, A, 200053941	2000-12-12
WO, A1, 9963083	1999-12-9	JP, A, 11332571	1999-12-7
		AU, A, 3955099	1999-12-20